

# HEAT KERNEL ESTIMATES FOR ANOMALOUS HEAVY-TAILED RANDOM WALKS

MATHAV MURUGAN AND LAURENT SALOFF-COSTE

**ABSTRACT.** Sub-Gaussian estimates for the natural random walk is typical of many regular fractal graphs. Subordination shows that there exist heavy tailed jump processes whose jump indices are greater than or equal to two. However, the existing machinery used to prove heat kernel bounds for such heavy tailed random walks fail in this case. In this work we extend Davies' perturbation method to obtain transition probability bounds for these anomalous heavy tailed random walks. We prove global upper and lower bounds on the transition probability density that are sharp up to constants. An important feature of our work is that the methods we develop are robust to small perturbations of the symmetric jump kernel.

## 1. INTRODUCTION

We are concerned with transition probability estimates for a class of random walks with heavy tailed jumps on fractal-like graphs. Many fractals satisfy the following sub-Gaussian estimates on the transition probability density  $p_t$  for the natural diffusion: there exists  $c_1, c_2, C_3, C_4 > 0$  such that

$$\begin{aligned} \frac{c_1}{t^{d_f/d_w}} \exp \left( - \left( \frac{d(x, y)^{d_w}}{c_2 t} \right)^{d_w/(d_w-1)} \right) \\ \leq p_t(x, y) \leq \frac{C_3}{t^{d_f/d_w}} \exp \left( - \left( \frac{d(x, y)^{d_w}}{C_4 t} \right)^{d_w/(d_w-1)} \right) \end{aligned} \quad (1)$$

for all points  $x, y$  in the underlying space and for all times  $t > 0$  [3, 12, 5, 28, 30]. The parameters  $d_f, d_w > 0$  are intrinsic and depend on the geometry of the underlying space. Sub-Gaussian estimates similar to (1) were obtained in the discrete time setting for the nearest neighbor walks on fractal like graphs [4, 29, 27, 26, 10, 6]. A precise formulation of these sub-Gaussian estimates for graphs is provided in [USG\( \$d\_f, d\_w\$ \)](#) and [LSG\( \$d\_f, d\_w\$ \)](#) below.

If  $\mathcal{L}$  denotes the generator of a diffusion whose heat kernel satisfies the sub-Gaussian estimate (1), then for any  $\beta \in (0, d_w)$ , the operator  $-(-\mathcal{L})^{\beta/d_w}$ , is the generator of a non-local Dirichlet form, and its heat kernel admits the estimate

$$p_t(x, y) \asymp \frac{1}{t^{d_f/\beta}} \left( 1 + \frac{d(x, y)}{t^{1/\beta}} \right)^{-(d_f+\beta)} \quad (2)$$

---

*Date:* December 9, 2015.

*2010 Mathematics Subject Classification.* 60J10, 60J75.

for all times  $t$  and for all points  $x, y$  in the underlying space (see, for example, [37], [31], [23, Lemma 5.4]). Here  $\asymp$  means that the ratio between both sides of the expression is bounded above and below by finite positive constants that do not depend on  $x, y, t$ . Estimates similar to (2) can be established using subordination in the discrete time context as well [35, Section 5]. Furthermore, it was shown in [25] that (1) and (2) exhaust all possible two-sided estimates of heat kernels of self-similar type. We call the parameter  $\beta$  in (2) the *jump index*.

Obtaining estimates of the form (2) for the case  $\beta \in (0, 2)$  has received much attention recently, both in the context of continuous time jump processes (see, for example, [16, 17, 9, 11]) and discrete time heavy tailed random walks (see [14, 34]). For the case  $\beta \in (0, 2)$ , there are methods to obtain estimates of the form (2) which are stable under bounded perturbations of the corresponding Dirichlet form. However it has been observed that the use of Davies' method to obtain the upper bound in (2) is no longer adequate for the case  $\beta \geq 2$  (see [24, Section 1], [35, Remark 1.4(d)]).

Recently characterizations of bounds of the form (2) for the case  $\beta \geq 2$  have been given in [24, Theorem 2.1 and Theorem 2.3]. However these characterizations rely on probabilistic estimates on the exit times (called survival estimates and tail estimates in [24]). It is not known if these survival and tail estimates are stable under bounded perturbation of the jump kernel of the jump process in the continuous time case or the conductance of the heavy tailed random walk in the discrete time case. Therefore there is a need to develop techniques to obtain (2) that are robust under bounded perturbation of the Dirichlet form for the case  $\beta \geq 2$ .

The main goal of the work is to show that the estimate (2) is stable under bounded perturbation of conductances for a heavy tailed random walk in the case  $\beta \geq 2$ , under some natural hypothesis. This is carried out by modifying the Davies' perturbation method using a cutoff Sobolev inequality. We extend the techniques developed in [36] to a non-local setting. Although our work concerns discrete time heavy tailed random walks, we expect that the techniques we develop extends to continuous time jump processes as well. For instance, in the proof of upper bound (see Proposition 3.1), we obtain corresponding continuous time bounds.

To state the results precisely, we introduce the general setup of a weighted graph. Let  $X$  be an infinite, connected, locally finite graph endowed with a *weight*  $\mu_{xy}$ . The elements of the set  $X$  are called vertices. Some of the vertices are connected by an edge, in which case we say that they are neighbors. The *weight* is a symmetric non-negative function on  $X \times X$  such that  $\mu_{xy} > 0$  if and only if  $x$  and  $y$  are neighbors (in which case we write  $x \sim y$ ). We call the pair  $(X, \mu)$  a *weighted graph*. The weight  $\mu_{xy}$  on the edges induces a weight  $\mu(x)$  on the vertices and a measure  $\mu$  on subsets  $A \subset X$  defined by

$$\mu(x) := \sum_{y: y \sim x} \mu_{xy} \quad \text{and} \quad \mu(A) := \sum_{x \in A} \mu(x).$$

Let  $d(x, y)$  be the graph distance between points  $x, y \in X$ , that is the minimal number of edges in any edge path connecting  $x$  and  $y$ . Denote the metric balls

and their measures as follows

$$B(x, r) := \{y \in X : d(x, y) \leq r\} \quad \text{and} \quad V(x, r) := \mu(B(x, r))$$

for all  $x \in X$  and  $r \geq 0$ . We assume that the measure  $\mu$  is comparable to the counting measure in the following sense: There exists  $C_\mu \in [1, \infty)$  such that  $\mu_x = \mu(\{x\})$  satisfies

$$C_\mu^{-1} \leq \mu_x \leq C_\mu \quad (3)$$

We consider weighted graphs  $(X, \mu)$  that satisfy the polynomial volume growth assumption: there exists  $d_f > 0$ ,  $C_V > 0$  such that

$$C_V^{-1} r^{d_f} \leq V(x, r) \leq C_V r^{d_f} \quad V(d_f)$$

for all  $r \geq 1$  and for all  $x \in X$ .

There is a natural random walk  $(X_n)_{n \in \mathbb{N}}$  on  $(X, \mu)$  associated with the edge weights  $\mu_{xy}$ . The Markov chain is defined by the following one-step transition probability

$$\mathbb{P}_x(X_1 = y) = \frac{\mu_{xy}}{\mu(x)}.$$

We will assume that there exists  $p_0 > 0$  such that  $P(x, y) \geq p_0$  for all  $x, y$  such that  $x \sim y$ . For any non-negative integer  $n$ , the  $n$ -step transition probability  $P_n$  is defined by  $P_n(x, y) = \mathbb{P}(X_n = y \mid X_0 = x) = \mathbb{P}^x(X_n = y)$ . Define the *heat kernel* of weighted graph  $(X, \mu)$  by

$$p_n(x, y) := \frac{P_n(x, y)}{\mu(y)}.$$

The Markov chain is symmetric with respect to measure  $\mu$ , that is  $p_n(x, y) = p_n(y, x)$  for all  $x, y \in X$  and all  $n \in \mathbb{N}$ . We assume that there exists  $d_w > 2$  such that the following *sub-Gaussian* estimates are true for the heat kernel  $p_n$ . There exists constants  $c, C > 0$  such that, for all  $x, y \in X$

$$p_n(x, y) \leq \frac{C}{n^{d_f/d_w}} \exp \left[ - \left( \frac{d(x, y)^{d_w}}{Cn} \right)^{\frac{1}{d_w-1}} \right], \forall n \geq 1 \quad \text{USG}(d_f, d_w)$$

and

$$(p_n + p_{n+1})(x, y) \geq \frac{c}{n^{d_f/d_w}} \exp \left[ - \left( \frac{d(x, y)^{d_w}}{cn} \right)^{\frac{1}{d_w-1}} \right], \forall n \geq 1 \vee d(x, y). \quad \text{LSG}(d_f, d_w)$$

The parameter  $d_w$  is called the *escape time exponent* or the *walk dimension* and the parameter  $d_f \geq 1$  is called the *volume growth exponent* or *fractal dimension* of the weighted graph  $(X, \mu)$ . It is known that  $d_w \geq 2$  necessarily (see for instance [19, Theorem 4.6]). For any  $d_f \in [1, \infty)$  and for any  $d_w \in [2, d_f + 1]$ , Barlow constructs graphs of polynomial volume growth satisfying  $V(x, r) \asymp (1 + r)^\alpha$  and sub-Gaussian estimates  $\text{USG}(d_f, d_w)$  and  $\text{LSG}(d_f, d_w)$  (see [4, Theorem 2] and [27, Theorem 3.1]). Moreover, these are the complete range of  $d_f$  and  $d_w$  for which sub-Gaussian estimates with escape rate exponent  $d_w$  could possibly hold for graphs of polynomial growth with volume growth exponent  $d_f$ .

We consider heavy tailed random walks on weighted graphs satisfying the above sub-Gaussian estimates. We say that a Markov operator  $K$  on a weighted graph  $(X, \mu)$  with volume growth exponent  $d_f$  satisfies  $\mathbf{J}(\beta)$ , if  $K$  has symmetric kernel  $k$  with respect to the measure  $\mu$  and if there exists a constant  $C > 0$  such that

$$C^{-1} \frac{1}{(1 + d(x, y))^{d_f + \beta}} \leq k(x, y) = k(y, x) \leq C \frac{1}{(1 + d(x, y))^{d_f + \beta}} \quad \mathbf{J}(\beta)$$

for all  $x, y \in \Gamma$ , where  $d_f$  is the volume group exponent given in  $\mathbf{V}(d_f)$ . If  $K$  satisfies  $\mathbf{J}(\beta)$  for some  $\beta > 0$ , then we say that  $\beta$  is the *jump index* of the random walk driven by  $K$ .

Let  $k_n(x, y)$  denote the kernel of the iterated power  $K^n$  with respect to the measure  $\mu$ . We are interested in obtaining estimates on  $k_n(x, y)$  for all values of  $n \in \mathbb{N}^*$  and for all  $x, y \in \Gamma$ . We say that a  $\mu$ -symmetric Markov operator  $K$  satisfies  $\mathbf{HKP}(d_f, \beta)$ , if there is a constant  $C > 0$  such that the iterated kernel  $k_n$  satisfies the estimate

$$C^{-1} \min \left( \frac{1}{n^{d_f/\beta}}, \frac{n}{(d(x, y))^{d_f + \beta}} \right) \leq k_n(x, y) \leq C \min \left( \frac{1}{n^{d_f/\beta}}, \frac{n}{(d(x, y))^{d_f + \beta}} \right) \quad \mathbf{HKP}(d_f, \beta)$$

for all  $x, y \in X$  and for all  $n \in \mathbb{N}^*$ , where  $d_f$  is the volume group exponent given by  $\mathbf{V}(d_f)$ . We say that  $K$  satisfies  $\mathbf{UHKP}(d_f, \beta)$ , if the iterated kernel  $k_n$  satisfies the upper bound in  $\mathbf{HKP}(d_f, \beta)$ . Similarly, we say that  $K$  satisfies  $\mathbf{LHKP}(d_f, \beta)$ , if the iterated kernel  $k_n$  satisfies the lower bound in  $\mathbf{HKP}(d_f, \beta)$ . Note that  $\mathbf{HKP}(d_f, \beta)$  is same as the bounds described in (2), since

$$\begin{aligned} \frac{1}{n^{d_f/\beta}} \left( 1 + \frac{d(x, y)}{n^{1/\beta}} \right)^{-(d_f + \beta)} &\leq \min \left( \frac{1}{n^{d_f/\beta}}, \frac{n}{(d(x, y))^{d_f + \beta}} \right) \\ \min \left( \frac{1}{n^{d_f/\beta}}, \frac{n}{(d(x, y))^{d_f + \beta}} \right) &\leq 2^{d_f + \beta} \frac{1}{n^{d_f/\beta}} \left( 1 + \frac{d(x, y)}{n^{1/\beta}} \right). \end{aligned}$$

The goal of this work is to develop methods to obtain the bound  $\mathbf{HKP}(d_f, \beta)$  that are robust to small perturbations of Dirichlet form in the sense given by  $\mathbf{J}(\beta)$ . Similar to the anomalous diffusion setting [36], we rely on a cutoff Sobolev inequality to implement Davies' method. To introduce cutoff Sobolev inequality, we first need to define cutoff functions and energy measure.

**Definition 1.1** (Cutoff function). Let  $U \subset V$  be finite sets in  $X$ . We say that  $\phi : X \rightarrow \mathbb{R}$  is a cutoff function for  $U \subset V$  if  $\phi \equiv 1$  on  $U$  and  $\phi \equiv 0$  on  $V^c$ .

**Definition 1.2** (Energy measure). For a  $\mu$ -symmetric Markov operator  $K$  with kernel  $k(x, y)$  with respect to  $\mu$ , and for functions  $f, g \in L^\infty(X, \mu)$ , we define the energy measure corresponding to  $K$  as the function

$$\Gamma(f, g)(x) = \frac{1}{2} \sum_{y \in X} (f(x) - f(y))(g(x) - g(y))k(x, y)\mu(y)\mu(x)$$

for all  $x \in X$ . The function  $\Gamma(f, g)$  can be considered as a (signed) measure where the measure of the singleton  $\{x\}$  is  $\Gamma(f, g)(x)$ .

We introduce a cutoff Sobolev inequality that plays an important role in obtaining  $\text{HKP}(d_f, \beta)$ . Cutoff Sobolev inequalities were first introduced in [7] for graphs and then extended to [9] for metric measure spaces. Recently Andres and Barlow simplified the cutoff Sobolev inequalities in [1]. We obtain transition probability estimates using the following cutoff Sobolev inequality. The motivation behind studying cutoff Sobolev inequalities in [7, 9, 1] is that they provide a method to obtain sub-Gaussian estimates that is robust to bounded perturbation of the Dirichlet form. We are motivated by similar reasons to formulate a version of cutoff Sobolev inequality for heavy tailed random walks. Our definition below is inspired by the cutoff Sobolev annulus inequality in [1] and a self-improving property of cutoff Sobolev inequality [1, Lemma 5.1].

**Definition 1.3** (Cutoff Sobolev inequality). Let  $K$  be a  $\mu$ -symmetric Markov operator on a weighted graph  $(X, \mu)$  and let  $\Gamma$  denote the corresponding energy measure. Let  $\beta \geq 2$ . We say that  $K$  satisfies the cutoff Sobolev inequality  $\text{CSJ}(\beta)$  if there exists  $C_1, C_2 > 0$  such that for all  $n \in \mathbb{N}^*$ , for all  $x \in X$ , for all  $r \geq 1$ , there exists a cut-off function  $\phi_n$  for  $B(x, r) \subset B(x, 2r)$  such that

$$\sum_{y \in X} f^2(y) \Gamma(\phi_n, \phi_n)(y) \leq \frac{C_1}{n^2} \sum_{y \in X} \Gamma(f, f)(y) + \frac{C_2 \mathcal{G}_\beta(n)}{r^\beta} \sum_{y \in X} f^2(y) \mu(y), \quad (4)$$

where the function  $\mathcal{G}_\beta$  is given by

$$\mathcal{G}_\beta(n) = \begin{cases} n^{\beta-2} & \text{if } \beta > 2 \\ \log(1+n) & \text{if } \beta = 2. \end{cases} \quad (5)$$

Further the function  $\phi_n$  above satisfies

$$\|\phi_n - \psi\|_\infty \leq n^{-1}, \quad (6)$$

where  $\psi$  is the linear cutoff function given by

$$\psi(y) = \min \left( \max \left( 1, \frac{2r - d(x, y)}{r} \right), 0 \right).$$

*Remark.* The parameter  $r$  in (4) in Definition 1.3 above is the primary parameter. The appearance of the expression  $r^\beta$  in the second term signifies the role played by  $\beta$  in the space-time scaling relation of the walk. The parameter  $n$  is auxiliary. In fact, it suffices to show the above cutoff Sobolev inequality (4) for the case  $n = 1$  due to a self-improving property. However, we will require (4) for arbitrarily large values of  $n$  in our proof of heat kernel bounds in Section 3.

The following transition probability estimate is the main result of our work.

**Theorem 1.4.** *Let  $(X, \mu)$  be a weighted graph satisfying (3) and  $\mathbf{V}(d_f)$  with volume growth exponent  $d_f$ . Suppose that the heat kernel of  $(X, \mu)$  satisfies the sub-Gaussian estimates  $\text{USG}(d_f, d_w)$  and  $\text{LSG}(d_f, d_w)$  with walk dimension  $d_w$ . Let  $K$  be a  $\mu$ -symmetric Markov operator whose kernel  $k = k_1$  satisfies  $\mathbf{J}(\beta)$  for some  $\beta \in [2, d_w)$ . Then the corresponding iterated kernel  $k_n$  satisfies  $\text{HKP}(d_f, \beta)$ .*

- Remark.* (a) For the case  $\beta \in (0, 2)$ ,  $\text{HKP}(d_f, \beta)$  follows from (3),  $\text{V}(d_f)$  and  $\text{J}(\beta)$  (see [34, Theorem 1.1]). In this case there is no assumption required on the heat kernel.
- (b) For the case  $\beta \geq d_w$ ,  $\text{HKP}(d_f, \beta)$  does not hold. We refer the reader to [35, Theorem 1.2] for on-diagonal estimates. We do not know off-diagonal estimates in this case.
- (c) We obtained matching two sided estimates on  $\sup_{x \in X} p_{2n}(x, x)$  in [35] for all  $\beta > 0$ . Due to Theorem 1.4, we have analogous pointwise on-diagonal lower bounds on  $k_n(x, x)$  for all  $x \in X$  for the case  $\beta \in [2, d_w)$ .
- (d) The conclusion of Theorem 1.4 holds if  $K$  is symmetric with respect to a different measure  $\nu$  that is comparable to the counting measure in the sense described by (3). The proof is robust to such perturbations of the symmetric measure.

The proof of Theorem 1.4 follows from the following stability result.

**Theorem 1.5** (Stability of  $\text{HKP}(d_f, \beta)$ ). *Let  $(X, \mu)$  be a weighted graph satisfying (3) and  $\text{V}(d_f)$  with volume growth exponent  $d_f$ . Let  $K_1$  and  $K_2$  be two  $\mu$ -symmetric Markov operators whose kernels with respect to  $\mu$  satisfy  $\text{J}(\beta)$  for some  $\beta \geq 2$ . Then  $K_1$  satisfies  $\text{HKP}(d_f, \beta)$  if and only if  $K_2$  satisfies  $\text{HKP}(d_f, \beta)$ .*

The proof of Theorem 1.4 using Theorem 1.5 is given below.

*Proof of Theorem 1.4.* By a known subordination argument (see [35, Theorem 5.1]), there exists a  $\mu$ -symmetric Markov operator satisfying  $\text{J}(\beta)$  and  $\text{HKP}(d_f, \beta)$ . The desired result then follows from Theorem 1.5.  $\square$

The rest of the paper is devoted to the proof of Theorem 1.5. The stability result in Theorem 1.5 is a consequence of the following characterization of  $\text{UHKP}(d_f, \beta)$  using  $\text{CSJ}(\beta)$ .

**Theorem 1.6.** *Let  $(X, \mu)$  be a weighted graph satisfying (3) and  $\text{V}(d_f)$  with volume growth exponent  $d_f$ . Let  $K$  be  $\mu$ -symmetric Markov operators whose kernel with respect to  $\mu$  satisfies  $\text{J}(\beta)$  for some  $\beta \geq 2$ . Then  $K$  satisfies  $\text{UHKP}(d_f, \beta)$  if and only if  $K$  satisfies  $\text{CSJ}(\beta)$ .*

## 2. CUTOFF SOBOLEV INEQUALITIES FOR HEAVY-TAILED RANDOM WALKS

The goal of this section is to prove the implication  $\text{UHKP}(d_f, \beta) \implies \text{CSJ}(\beta)$  in Theorem 1.6.

**Proposition 2.1.** *Let  $(X, \mu)$  be a weighted graph satisfying (3),  $\text{V}(d_f)$  with volume growth exponent  $d_f$ . Let  $K$  be  $\mu$ -symmetric Markov operator whose kernel with respect to  $\mu$  satisfies  $\text{J}(\beta)$  for some  $\beta \geq 2$ . If  $K$  satisfies  $\text{UHKP}(d_f, \beta)$ , then  $K$  satisfies the cutoff Sobolev inequality  $\text{CSJ}(\beta)$ .*

We follow the approach of Andres and Barlow in [1, Section 5] to prove the cutoff Sobolev inequality in Proposition 2.1. The first step is to obtain estimates on exit times. We define the exit time for the Markov chain  $(Y_n)_{n \in \mathbb{N}}$  as

$$\tau_{B(x,r)}^Y = \min \{k \in \mathbb{N} : Y_k \notin B(x, r)\}.$$

The exit time  $\tau_{B(x,r)}^Y$  satisfies the following survival estimate.

**Lemma 2.2.** *Let  $(X, \mu)$  be a weighted graph satisfying (3),  $\mathbf{V}(d_f)$  with volume growth exponent  $d_f$ . Let  $K$  be  $\mu$ -symmetric Markov operator satisfying  $\mathbf{UHKP}(d_f, \beta)$ , for some  $\beta \geq 2$ . Let  $(Y_n)_{n \in \mathbb{N}}$  denote the Markov chain driven by the operator  $K$ . There exists constants  $\epsilon, \delta \in (0, 1)$  such that for all  $x \in X$  and for all  $r \geq 1$ ,*

$$\mathbb{P}_x \left( \tau_{B(x,r)}^Y \leq \delta r^\beta \right) \leq \epsilon. \quad (7)$$

*Proof.* We follow the argument in [11, pp. 15]. By  $\mathbf{UHKP}(d_f, \beta)$  and  $\mathbf{V}(d_f)$ , there exists  $C_1, C_2 \geq 1$  such that

$$\mathbb{P}_x(d(Y_n, x) \geq r) = \sum_{y \notin B(x,r)} k_n(x, y) \mu(y) \leq C_1 \sum_{y \in B(x,r)^c} \frac{n \mu(y)}{d(x, y)^{d_f + \beta}} \leq C_2 \frac{n}{r^\beta} \quad (8)$$

for all  $n \in \mathbb{N}^*$  and for all  $x \in X$ . By (8) and the strong Markov property of  $\{Y_k\}$  at time  $\tau = \tau_{B(x,r)}^Y$ , there exists  $C_3 \geq 1$  such that

$$\begin{aligned} \mathbb{P}_x(\tau \leq n) &\leq \mathbb{P}_x(\tau \leq n, d(Y_{2n}, x) \leq r/2) + \mathbb{P}_x(d(Y_{2n}, x) > r/2) \\ &\leq \mathbb{P}_x(\tau \leq n, d(Y_{2n}, Y_\tau) \geq r/2) + 2^{1+\beta} C_2 n / r^\beta \\ &= \mathbb{E}^x(\mathbf{1}_{\tau \leq n} \mathbb{P}_{Y_\tau}(d(Y_{2n-\tau}, Y_0) \geq r/2)) + 2^{1+\beta} C_2 n / r^\beta \\ &\leq \sup_{y \in B(x,r)^c} \sup_{s \leq n} \mathbb{P}^y(d(Y_{2n-s}, y) \geq r/2) + 2^{1+\beta} C_2 n / r^\beta \\ &\leq C_3 n / r^\beta \end{aligned}$$

for all  $x \in X$ ,  $k \in \mathbb{N}^*$  and for all  $r > 0$ . This immediately implies the desired bound (7).  $\square$

For  $D \subset X$  and a  $\mu$ -symmetric Markov operator  $K$  with kernel  $k$ , we define the sub-Markov operator  $K_D$  corresponding to the walk killed upon exiting  $D$  defined by

$$K_D f(x) = \sum_{y \in D} f(y) k_D(x, y) \mu(y)$$

for all  $x \in X$ , where  $k_D(x, y) = k(x, y) \mathbf{1}_D(x) \mathbf{1}_D(y)$ . As before, we define the exit time of  $D$  as

$$\tau_D = \tau_D^Y = \min \{n \in \mathbb{N} : Y_n \notin D\},$$

where  $(Y_n)_{n \in \mathbb{N}}$  is the Markov chain corresponding to the operator  $K$ . For  $D \subset X$ ,  $\lambda > 0$ , we define corresponding the ‘resolvent operator’ as

$$G_\lambda^D f(x) = \left( I - \frac{K_D}{1 + \lambda} \right)^{-1} f(x) = \sum_{i=0}^{\infty} (1 + \lambda)^{-i} K_D^i f(x) = \mathbb{E}_x \sum_{i=0}^{\tau_D} (1 + \lambda)^{-i} f(Y_i). \quad (9)$$

**Lemma 2.3.** *Let  $(X, \mu)$  be a weighted graph satisfying (3) and  $\mathbf{V}(d_f)$  with volume growth exponent  $d_f$ . Let  $K$  be a Markov operator satisfying  $\mathbf{UHKP}(d_f, \beta)$ . Let  $x_0 \in X$ ,  $r > 10$ ,  $R > 0$  and define the annuli  $D_0 = B(x_0, R + 9r/10) \setminus B(x_0, R +$*

$r/10$ ),  $D_1 = B(x_0, R + 4r/5) \setminus B(x_0, R + r/5)$ ,  $D_2 = B(x_0, R + 3r/5) \setminus B(x_0, R + r/5)$ . Let  $\lambda = r^{-\beta}$  and set

$$h = G_\lambda^{D_0} \mathbf{1}_{D_1}, \quad (10)$$

where  $G_\lambda^{D_0}$  is as defined in (9). Then  $h$  is supported in  $D_0$  and satisfies

$$h(x) \leq 2r^\beta \quad \text{for all } x \in X, \quad (11)$$

$$h(x) \geq c_1 r^\beta \quad \text{for all } x \in D_2, \quad (12)$$

*Proof.* Since  $K_{D_0}$  is a contraction in  $L^\infty$ , we have

$$h(x) = \sum_{i=0}^{\infty} (1 + \lambda)^{-i} K_{D_0}^i \mathbf{1}_{D_1}(x) \leq \sum_{i=0}^{\infty} (1 + \lambda)^{-i} = (1 + \lambda)^{-1} \leq 2r^\beta$$

for all  $r \geq 1$ , for all  $x_0, x \in X$ .

Let  $(Y_n)_{n \in \mathbb{N}}$  denote the Markov chain driven by  $K$ . Let  $\epsilon, \delta \in (0, 1)$  be given by Lemma 2.2. Let  $r_0 = r/5$ ,  $x \in D_2$  and  $B_1 = B(x, r_0) \subset D_1$ . Then there exists  $c_1 > 0$  such that for all  $x_0 \in X$ , for all  $r > 10$ , for all  $x \in D_2$ , we have

$$\begin{aligned} h(x) &= \sum_{i=0}^{\infty} (1 + \lambda)^{-i} K_{D_0}^i \mathbf{1}_{D_1}(x) \geq \sum_{i=0}^{\infty} (1 + \lambda)^{-i} K_{B_1}^i \mathbf{1}_{B_1}(x) \\ &\geq (1 + r^{-\beta})^{-\delta r^\beta} \mathbb{E}_x \left[ \sum_{i=0}^{\lfloor \delta r^\beta \rfloor \wedge \tau_{B_1}^Y} \mathbf{1}_{B_1}(Y_i) \right] \\ &\geq (1 + r^{-\beta})^{-\delta r^\beta} \delta r^\beta \mathbb{P}_x(\tau_{B_1}^Y > \delta r^\beta) \geq (1 + r^{-\beta})^{-\delta r^\beta} \delta r^\beta (1 - \epsilon) \geq c_1 r^\beta. \end{aligned}$$

□

The Dirichlet form corresponding to the operator  $P$  is defined as

$$\mathcal{E}_P(f, g) := \langle f, (I - P)g \rangle$$

for all  $f, g \in L^2(X, \mu)$ , where  $\langle \cdot, \cdot \rangle$  denotes the inner product in  $L^2(X, \mu)$ .

**Lemma 2.4** (Folklore). *Let  $P$  be a  $\mu$ -symmetric sub-Markov operator with kernel  $p$  and let  $\mathcal{E}_P$  and  $\Gamma$  denote the corresponding Dirichlet form and energy measure respectively.*

(a) *We have*

$$\mathcal{E}_P(f, g) = \sum_{x \in X} \Gamma(f, g) + \sum_{x \in X} f(x)g(x)(1 - P\mathbf{1}(x))\mu(x) \quad (13)$$

for all  $f, g \in L^2(X, \mu)$ . In particular, if  $P$  is a Markov operator we have

$$\mathcal{E}_P(f, g) = \langle f, (I - P)g \rangle = \sum_{x \in X} \Gamma(f, g). \quad (14)$$

(b) *The energy measure  $\Gamma$  satisfies the integrated version of Leibnitz rule*

$$\sum_{x \in X} \Gamma(fg, h)(x) = \sum_{x \in X} [f(x)\Gamma(g, h)(x) + g(x)\Gamma(f, h)(x)] \quad (15)$$



for all bounded functions  $f, g, h$ , as long as the above sums are absolutely convergent.

*Proof.* Note that

$$\begin{aligned}\mathcal{E}_P(f, g) &= \langle f, (I - P)g \rangle = \sum_{x \in X} f(x)g(x)\mu(x) - \sum_{x \in X} f(x)Pg(x)\mu(x) \\ &= \sum_{x \in X} f(x)g(x)\mu(x)(1 - P\mathbf{1}(x) + P\mathbf{1}(x)) - \sum_{x \in X} f(x)Pg(x)\mu(x) \quad (16)\end{aligned}$$

By symmetry  $p(x, y) = p(y, x)$ , we have

$$\begin{aligned}\sum_{x \in X} \Gamma(f, g)(x) &= \frac{1}{2} \sum_{x, y} (f(x) - f(y))(g(x) - g(y))p(x, y)\mu(x)\mu(y) \\ &= \frac{1}{2} \sum_{x \in X} f(x)g(x)P\mathbf{1}(x)\mu(x) + \frac{1}{2} \sum_{y \in X} f(y)g(y)P\mathbf{1}(y)\mu(y) \\ &\quad - \frac{1}{2} \sum_{x \in X} f(x)Pg(x)\mu(x) - \frac{1}{2} \sum_{y \in X} f(y)Pg(y)\mu(y) \\ &= \sum_{x \in X} f(x)g(x)P\mathbf{1}(x)\mu(x) - \sum_{x \in X} f(x)Pg(x)\mu(x) \quad (17)\end{aligned}$$

Combining (16) and (17), we obtain (13).

To prove (b), we follow [15, Theorem 3.7] and write

$$\begin{aligned}&(f(x)g(x) - f(y)g(y))(h(x) - h(y)) \\ &= \frac{1}{2}(g(x) + g(y))(f(x) - f(y))(h(x) - h(y)) \\ &\quad + \frac{1}{2}(f(x) + f(y))(g(x) - g(y))(h(x) - h(y)).\end{aligned}$$

Then an application of symmetry  $p(x, y) = p(y, x)$  similar to (a) yields the desired result (15).  $\square$

The following technical lemma is a consequence of Leibnitz rule above.

**Lemma 2.5.** ([15, Lemma 3.5]) *Let  $(X, \mu)$  be a weighted graph and let  $P$  with a  $\mu$ -symmetric sub-Markov operator with kernel  $p$  with respect to  $\mu$ . Let  $\Gamma$  denote the corresponding energy measure. If  $f, h \in L^2(X, \mu)$  and  $g \in L^\infty(X, \mu)$ , we have*

$$\sum_{x \in X} g(x)\Gamma(f, h)(x) = \frac{1}{2} \sum_{x \in X} [\Gamma(gh, f)(x) + \Gamma(gf, h)(x) - \Gamma(g, fh)(x)].$$

*Proof.* We use Leibnitz rule (Lemma 2.4(b)) to all terms in the right hand side to obtain the desired equality. The absolute convergence of the various sums are a consequence of Hölder inequality.  $\square$

**Proposition 2.6.** *Let  $(X, \mu)$  be a weighted graph satisfying (3) and  $\mathbf{V}(d_f)$ . Let  $K$  be a  $\mu$ -symmetric Markov operator satisfying  $\mathbf{UHKP}(d_f, \beta)$ , for some  $\beta \geq 2$ . Then there exists  $C_1, C_2 > 0$  satisfying the following properties: for any  $x \in X$*

and for any  $R, r > 0$ , there exists a cutoff function  $\phi$  for  $B(x, R) \subset B(x, R + r)$  such that

$$\phi \equiv 1 \quad \text{in } B(x, R + r/2), \quad (18)$$

$$\phi \equiv 0 \quad \text{in } B(x, R + 9r/10)^c, \quad (19)$$

and for any function  $f \in L^2(X, \mu)$

$$\sum_{y \in U} f^2(y) \Gamma_U(\phi, \phi)(y) \leq C_1 \sum_{y \in U} \Gamma_U(f, f)(y) + \frac{C_2}{r^\beta} \sum_{y \in U} f^2(y) \mu(y), \quad (20)$$

where  $U = B(x, R + r) \setminus B(x, R)$  and  $\Gamma_U$  is the energy measure corresponding to the sub-Markov operator  $K_U$ .

*Proof.* If  $r \leq 10$ , we may assume  $\phi = \mathbf{1}_{B(x, R+r/2)}$ . Let  $r > 10$  and let  $h, c_1$  be as defined in Lemma 2.3. Set

$$g(y) = \frac{h(y)}{c_1 r^\beta} \quad \text{for all } y \in X \quad (21)$$

$$\phi(y) = \begin{cases} 1 \wedge g(y) & \text{if } y \in B(x, R + r/2)^c \\ 1 & \text{if } y \in B(x, R + r/2) \end{cases} \quad (22)$$

By Lemma 2.3, we have (18) and (19). It remains to verify (20). By Leibnitz rule (Lemma 2.4(b), (15)) and the fact that  $\Gamma_U(g, g)$  is supported in  $U$ , we obtain

$$\begin{aligned} \sum_{y \in U} f^2(y) \Gamma_U(g, g)(y) &= \sum_{y \in X} f^2(y) \Gamma_U(g, g)(y) \\ &= \sum_{y \in X} \Gamma_U(f^2 g, g)(y) - \sum_{y \in X} g(y) \Gamma_U(f^2, g)(y). \end{aligned} \quad (23)$$

As in the Lemma 2.3, we set  $\lambda = r^{-\beta}$  and  $D_1 = B(x_0, R + 4r/5) \setminus B(x_0, R + r/5)$ . Since  $g$  is supported in  $D_0 = B(x, R + 9r/10) \setminus B(x, R + r/10)$ , by (13) of Lemma 2.4(a) and Lemma 2.3, we obtain

$$\begin{aligned} \sum_{y \in X} \Gamma_U(f^2 g, g) &\leq \mathcal{E}_{K_U}(f^2 g, g) \leq \langle f^2 g, (I - K_U)g \rangle_{L^2(U, \mu)} + \lambda \langle f^2 g, g \rangle_{L^2(U, \mu)} \\ &= \langle f^2 g, ((1 + \lambda)I - K_U)g \rangle_{L^2(U, \mu)} \\ &= \langle f^2 g, ((1 + \lambda)I - K_{D_0})g \rangle_{L^2(D_0, \mu)} \\ &= (1 + \lambda)(c_1 r^\beta)^{-1} \langle f^2 g, (I - (1 + \lambda)^{-1} K_U) G_\lambda^{D_0} \mathbf{1}_{D_1} \rangle_{L^2(D_0, \mu)} \\ &= (1 + \lambda)(c_1 r^\beta)^{-1} \langle f^2 g, \mathbf{1}_{D_1} \rangle_{L^2(D_0, \mu)} \\ &\leq 4c_1^{-2} r^{-\beta} \sum_{y \in D_1} f^2(y) \mu(y). \end{aligned} \quad (24)$$

We use Cauchy-Schwarz inequality, the  $\mu$ -symmetry of  $K_U$  and  $ab \leq a^2/4 + b^2$  to obtain

$$\begin{aligned}
& \left| \sum_{y \in X} g(y) \Gamma_U(f^2, g)(y) \right| \\
&= \frac{1}{2} \left| \sum_{y, z \in U} g(y) (f^2(y) - f^2(z)) (g(y) - g(z)) k(x, y) \mu(y) \mu(z) \right| \\
&\leq \frac{1}{2} \left| \sum_{y, z \in U} g(y) f(y) (f(y) - f(z)) (g(y) - g(z)) k(x, y) \mu(y) \mu(z) \right| \\
&\quad + \frac{1}{2} \left| \sum_{y, z \in U} g(y) f(z) (f(y) - f(z)) (g(y) - g(z)) k(x, y) \mu(y) \mu(z) \right| \\
&\leq \frac{1}{4} \sum_{y \in U} f^2(y) \Gamma_U(g, g)(y) + \sum_{y \in U} g^2(y) \Gamma_U(f, f)(y) \\
&\quad + \frac{1}{4} \sum_{z \in U} f^2(z) \Gamma_U(g, g)(z) + \sum_{y \in U} g^2(y) \Gamma_U(f, f)(y) \\
&= \frac{1}{2} \sum_{y \in U} f^2(y) \Gamma_U(g, g)(y) + 2 \sum_{y \in U} g^2(y) \Gamma_U(f, f)(y) \tag{25}
\end{aligned}$$

Combining the above, we obtain

$$\begin{aligned}
\sum_{y \in U} f^2(y) \Gamma(\phi, \phi)(y) &\leq \sum_{y \in U} f^2(y) \Gamma(g, g)(y) \\
&\leq 2 \sum_{y \in X} \Gamma_U(f^2 g, g)(y) + 4 \sum_{y \in U} g^2(y) \Gamma_U(f, f)(y) \\
&\leq 4c_1^{-2} \sum_{y \in U} g^2(y) \Gamma_U(f, f)(y) + 8c_1^{-2} r^{-\beta} \sum_{y \in U} f^2(y) \mu(y).
\end{aligned}$$

For the first line above we use  $|\phi(y) - \phi(z)| \leq |g(y) - g(z)|$ , the second line follows from (23) and (25), and the last line follows from (24) and  $g \leq c_1^{-1}$   $\square$

The “linear cutoff function” for  $B(x, r) \subset B(x, 2r)$

$$\psi(y) = \min \left( \max \left( 1, \frac{2r - d(x, y)}{r} \right), 0 \right) \tag{26}$$

is commonly used to obtain off-diagonal estimates using Davies’ method (see, for example, [21, 11]). We will first see how this linear cutoff functions compares to the ones obtained in Proposition 2.6 satisfying inequality (20).

**Lemma 2.7.** *Let  $(X, \mu)$  be a weighted graph satisfying (3) and  $\mathbf{V}(d_f)$ . Let  $K$  be a  $\mu$ -symmetric Markov operator satisfying  $\mathbf{UHKP}(d_f, \beta)$ , for some  $\beta \geq 2$ . Let  $\psi$  is the cutoff function in (26) for  $B(x, R) \subset B(x, R + r)$  for some  $x \in X$  and*

$r \geq 1$ . Then there exists  $C_1 > 0$  such that the corresponding energy measure  $\Gamma(\psi, \psi)$  satisfies the inequality

$$\Gamma(\psi, \psi)(y) \leq \begin{cases} \frac{C_1}{r^2} & \text{if } \beta > 2 \\ \frac{C_1 \log(1+r)}{r^2} & \text{if } \beta = 2 \end{cases} \quad (27)$$

for all  $y \in X$ , where  $C_1$  does not depend on  $x \in X$ ,  $r \geq 1$  or  $R > 0$ .

*Proof.* Let  $k$  denote the  $\mu$ -symmetric kernel of  $K$ . Note that  $\psi$  is  $1/r$ -Lipschitz function and  $|\psi(y) - \psi(z)| \leq 1$  for all  $y, z \in X$ . Therefore

$$\begin{aligned} \Gamma(\psi, \psi)(y) &= \sum_{z \in X} (\psi(y) - \psi(z))^2 k(y, z) \mu(z) \mu(y) \\ &\leq \sum_{z \in B(y, r)} r^{-2} d(y, z)^2 k(y, z) \mu(z) \mu(y) + \sum_{z \notin B(y, r)} k(y, z) \mu(z) \mu(y) \end{aligned} \quad (28)$$

To bound the second term above, by **J**( $\beta$ ) and **V**( $d_f$ ) there exists  $C_2, C_3 > 0$  such that

$$\begin{aligned} \sum_{z \notin B(y, r)} k(y, z) \mu(z) \mu(y) &\leq \sum_{z \notin B(y, r)} \frac{C_2 \mu(z)}{(d(y, z))^{d_f + \beta}} \\ &= \sum_{i=1}^{\infty} \sum_{z: 2^{i-1}r < d(y, z) \leq 2^i r} \frac{C_2 \mu(z)}{(d(y, z))^{d_f + \beta}} \leq \sum_{i=1}^{\infty} \frac{C_2 V(y, 2^i r)}{(2^{i-1}r)^{d_f + \beta}} \\ &\leq C_3 r^{-\beta} \leq C_3 r^{-2} \end{aligned} \quad (29)$$

for all  $y \in X$  and for all  $r \geq 1$ .

For the first term in (28), by **J**( $\beta$ ) and **V**( $d_f$ ) there exists  $C_4, C_5 > 0$  such that

$$\begin{aligned} \sum_{z \in B(y, r)} d(y, z)^2 k(y, z) \mu(z) \mu(y) &\leq \sum_{i=0}^{\lfloor \log_2 r \rfloor} \sum_{z \in X: 2^i \leq d(y, z) < 2^{i+1}} d(y, z)^2 k(y, z) \mu(z) \mu(y) \\ &\leq C_4 \sum_{i=0}^{\lfloor \log_2 r \rfloor} 2^{i(2-\beta)} \\ &\leq \begin{cases} C_5 \log(1+r) & \text{if } \beta = 2 \\ C_5 & \text{if } \beta > 2 \end{cases} \end{aligned} \quad (30)$$

for all  $y \in X$  and for all  $r \geq 1$ . Combining (28), (29) and (30), we obtain the desired bound (27).  $\square$

For a  $\mu$ -symmetric Markov operator  $K$  with kernel  $k(\cdot, \cdot)$  and for bounded function  $f$ , we define the energy measure corresponding to a truncation at scale  $L > 0$  as

$$\Gamma_{(L)}(f, f)(y) = \frac{1}{2} \sum_{y \in B(x, L)} (f(x) - f(y))^2 k(x, y) \mu(y) \mu(x). \quad (31)$$

*Proof of Proposition 2.1.* Let  $\Gamma$  denote the energy measure of  $K$ . By Lemma 2.7, without loss of generality we may assume  $r > 10n$ . We divide the annulus  $U = B(x, 2r) \setminus B(x, r)$  into  $n$ -annuli  $U_1, U_2, \dots, U_n$  of equal width, where

$$U_i := B(x, r + ir/n) \setminus B(x, r + (i-1)r/n), \quad i = 1, 2, \dots, n.$$

By Proposition 2.6, there exists  $C_3, C_4 > 0$  and cutoff functions  $\phi_{(i)}$  for  $B(x, R + (i-1)r/n) \subset B(x, R + ir/n)$  satisfying  $0 \leq \phi_{(i)} \leq 1$ ,

$$\phi_{(i)} \equiv 1 \quad \text{in } B(x, r + (i-1)r/n + r/(2n)), \quad (32)$$

$$\phi_{(i)} \equiv 0 \quad \text{in } B(x, r + (i-1)r/n + 9r/(10n))^c, \quad (33)$$

$$\sum_{y \in U_i} f^2(y) \Gamma_{U_i}(\phi_{(i)}, \phi_{(i)}) \leq C_3 \sum_{y \in U_i} \Gamma_{U_i}(f, f)(y) + \frac{C_4}{(r/n)^\beta} \sum_{U_i} f^2(y) \mu(y) \quad (34)$$

for  $i = 1, 2, \dots, n$ . We define  $\phi_n = n^{-1} \sum_{i=1}^n \phi_{(i)}$ ,  $s = r/(10n)$  and  $\Gamma_{(s)}$  truncated energy measure at scale  $s$  as given by (31). Note that  $\phi_n$  satisfies (6) because  $\phi_n(y) \in [1 - i/n, 1 - (i-1)/n]$  for all  $y \in U_i$  and for all  $i = 1, 2, \dots, n$ . By (32) and (33), we have

$$\Gamma_{(s)}(\phi_n, \phi_n)(y) = n^{-2} \sum_{i=1}^n \Gamma_{(s)}(\phi_{(i)}, \phi_{(i)})(y) \leq \sum_{i=1}^n \Gamma_{U_i}(\phi_{(i)}, \phi_{(i)})(y) \quad (35)$$

for all  $y \in X$ . Combining (34) and (35), we obtain

$$\sum_{y \in X} f^2(y) \Gamma_{(s)}(\phi_n, \phi_n)(y) \leq C_3 n^{-2} \sum_{y \in X} \Gamma(f, f)(y) + \frac{C_4 n^{\beta-2}}{r^\beta} \sum_{y \in X} f^2(y) \mu(y). \quad (36)$$

To bound  $\Gamma(\phi_n, \phi_n)(y) - \Gamma_{(s)}(\phi_n, \phi_n)(y)$ , we write

$$\Gamma(\phi_n, \phi_n) - \Gamma_{(s)}(\phi_n, \phi_n) = \Gamma(\phi_n, \phi_n) - \Gamma_{(r)}(\phi_n, \phi_n) + \Gamma_{(r)}(\phi_n, \phi_n) - \Gamma_{(s)}(\phi_n, \phi_n). \quad (37)$$

Since  $0 \leq \phi_n \leq 1$  by (29) there exists  $C_5, C_6 > 0$  such that

$$\Gamma(\phi_n, \phi_n)(y) - \Gamma_{(r)}(\phi_n, \phi_n)(y) \leq \sum_{y \in B(x, r)^c} k(x, y) \mu(x) \mu(y) \leq C_5 r^{-\beta} \leq C_6 r^{-\beta} \quad (38)$$

for all  $r > 0$  and  $n \in \mathbb{N}^*$ . Note that for all  $y, z \in X$  such that  $d(x, y) \geq s = r/(10n)$ , by (6) we have  $|\phi_n(y) - \phi_n(z)| \leq r^{-1} d(y, z) + 2/n \leq 21r^{-1} d(y, z)$ . Therefore by a similar calculation as (30), there exists  $C_7 > 0$  such that, for all  $y \in X$ , for all

$n \in \mathbb{N}^*$  and for all  $r > 10n$ , we have

$$\begin{aligned}
& \Gamma_{(r)}(\phi_n, \phi_n)(y) - \Gamma_{(s)}(\phi_n, \phi_n)(y) \\
&= \frac{1}{2} \sum_{z \in B(y, r) \setminus B(y, s)} (\phi_n(y) - \phi_n(z))^2 k(y, z) \mu(x) \mu(y) \\
&\leq \frac{21^2}{2} \sum_{z \in B(y, r) \setminus B(y, s)} r^{-2} d(y, z)^2 k(y, z) \mu(x) \mu(y) \\
&\leq \frac{21^2}{2} \sum_{i=0}^{\lfloor \log_2(10n) \rfloor} \sum_{z \in B(y, 2^{i+1}s) \setminus B(y, 2^i s)} r^{-2} d(y, z)^2 k(y, z) \mu(x) \mu(y) \\
&\leq C_7 \frac{\mathcal{G}_\beta(n)}{r^\beta}.
\end{aligned} \tag{39}$$

Combining (36), (37), (38) and (39), we obtain

$$\sum_{y \in X} f^2(y) \Gamma(\phi_n, \phi_n)(y) \leq \frac{C_3}{n^2} \mathcal{E}_K(f, f) + (C_4 + C_6 + C_7) \frac{\mathcal{G}_\beta(n)}{r^\beta} \sum_{y \in X} f^2(y) \mu(y) \tag{40}$$

for all  $x \in X$ , for all  $n \in \mathbb{N}^*$  and for all  $r > 10n$ . Combining (37) with Lemma 2.7 yields the desired result.  $\square$

### 3. DAVIES' METHOD

In this section, we carry out the Davies perturbation method to obtain heat kernel upper bounds  $\text{UHKP}(d_f, \beta)$  for heavy tailed walks satisfying  $\mathbf{J}(\beta)$  and the cutoff Sobolev inequality  $\text{CSJ}(\beta)$ , for some  $\beta \geq 2$ . For the case the  $\beta \in (0, 2)$ , in [11] the heat kernel upper bounds for the corresponding continuous time process was obtained. The corresponding discrete time bounds were obtained in [34].

The idea behind the approach of [11] is to use Meyer's construction [32] to split the jump kernel into small and large sums and then apply Davies' method for the smaller jumps (see [11, Section 3]). However as mentioned in the introduction (see [24, Section 1]), the above approach is no longer adequate to obtain  $\text{UHKP}(d_f, \beta)$  for the case  $\beta \geq 2$ . The goal of this section is to modify Davies' perturbation method to obtain upper bounds for heavy tailed jump processes satisfying  $\mathbf{J}(\beta)$  for the case  $\beta \geq 2$ . The following Proposition is the converse of Proposition 2.1.

**Proposition 3.1.** *Let  $(X, \mu)$  be a weighted graph satisfying (3),  $\mathbf{V}(d_f)$  with volume growth exponent  $d_f$ . Let  $K$  be  $\mu$ -symmetric Markov operator whose kernel with respect to  $\mu$  satisfies  $\mathbf{J}(\beta)$  for some  $\beta \geq 2$ . If  $K$  satisfies the cutoff Sobolev inequality  $\text{CSJ}(\beta)$ , then  $K$  satisfies the transition probability upper bounds  $\text{UHKP}(d_f, \beta)$ .*

We introduce two definitions below.

**Definition 3.2.** We say that a  $\mu$ -symmetric sub-Markov operator  $T$  on a weighted graph  $(X, \mu)$  is  $L$ -local if its corresponding kernel  $t$  satisfies  $t(x, y) = 0$  whenever  $x, y \in X$  satisfies  $d(x, y) > L$ .

**Definition 3.3.** Let  $\psi$  be a function on a weighted graph  $(X, \mu)$  and let  $L > 0$ . We define the oscillation of  $\psi$  at scale  $L$  as

$$\text{osc}(\psi, L) := \sup_{x, y \in X: d(x, y) \leq L} |\psi(x) - \psi(y)|.$$

The following Lemma is an analog of [15, Theorem 3.9]. The computations are similar to the ones in [15] but we will use a different strategy to control the energy measure at various places.

**Lemma 3.4.** *Let  $T$  be a  $\mu$ -symmetric,  $L$ -local, sub-Markov operator on a weighted graph  $(X, \mu)$  and let  $\Gamma$  denote the corresponding energy measure. Then for any function  $\psi \in L^\infty(X, \mu)$ , for all  $p \geq 1$  and for all  $f \in L^2(X, \mu)$  with  $f \geq 0$ , we have*

$$\begin{aligned} & \sum_{x \in X} \Gamma(e^{-\psi} f, e^{\psi} f^{2p-1})(x) \\ & \geq \frac{1}{2p} \sum_{x \in X} \Gamma(f^p, f^p)(x) - 9pe^{2\text{osc}(\psi, L)} \sum_{x \in X} f^{2p}(x) \Gamma(\psi, \psi)(x). \end{aligned} \quad (41)$$

*Proof.* Let  $t$  denote the kernel of  $T$  with respect to  $\mu$ . Then by Lemma 2.5, we have

$$\begin{aligned} \sum_{x \in X} \Gamma(e^{-\psi} f, e^{\psi} f^{2p-1})(x) &= \sum_{x \in X} \left( \Gamma(e^{-\psi} f^{2p}, e^{\psi})(x) + \Gamma(f, f^{2p-1})(x) \right) \\ &\quad - 2 \sum_{x \in X} e^{-\psi(x)} f(x) \Gamma(e^{\psi}, f^{2p-1})(x). \end{aligned} \quad (42)$$

We use Leibniz rule (Lemma 2.4(b)) for the first term in the right in (42) to obtain

$$\begin{aligned} \sum_{x \in X} \Gamma(e^{-\psi} f, e^{\psi} f^{2p-1})(x) &= \sum_{x \in X} f^{2p-1}(x) \Gamma(e^{-\psi} f, e^{\psi})(x) \\ &\quad + \sum_{x \in X} \left( \Gamma(f, f^{2p-1})(x) - e^{-\psi(x)} f(x) \Gamma(e^{\psi}, f^{2p-1})(x) \right). \end{aligned} \quad (43)$$

Note that

$$\begin{aligned} & \sum_{x \in X} \left( f^{2p-1}(x) \Gamma(e^{-\psi} f, e^{\psi})(x) - e^{-\psi(x)} f(x) \Gamma(e^{\psi}, f^{2p-1})(x) \right) \\ &= \frac{1}{2} \sum_{x, y \in X} \left\{ \left( e^{-\psi(x)} f(x) f^{2p-1}(y) - e^{-\psi(y)} f(y) f^{2p-1}(x) \right) \right. \\ &\quad \left. \times \left( e^{\psi(x)} - e^{\psi(y)} \right) t(x, y) \mu(x) \mu(y) \right\} \end{aligned} \quad (44)$$

We define  $a(x, y) := t(x, y)\mu(x)\mu(y)$  for all  $x, y \in X$ . Next, we rewrite the right side of (44) as

$$\begin{aligned} & \frac{1}{2} \sum_{x, y \in X} (f^{2p}(y) - f^{2p}(x))e^{-\psi(x)}(e^{\psi(x)} - e^{\psi(y)})a(x, y) \\ & + \frac{1}{2} \sum_{x, y \in X} f^{2p}(x)(e^{-\psi(x)} - e^{-\psi(y)})(e^{\psi(x)} - e^{\psi(y)})a(x, y) \\ & + \sum_{x, y \in X} f^{2p-1}(y)(f(x) - f(y))e^{\psi(y)}(e^{-\psi(y)} - e^{-\psi(x)})a(x, y) \end{aligned} \quad (45)$$

For the first term in (45), we use Cauchy-Schwartz inequality to obtain

$$\begin{aligned} & \frac{1}{2} \sum_{x, y \in X} (f^{2p}(y) - f^{2p}(x))e^{-\psi(x)}(e^{\psi(x)} - e^{\psi(y)})a(x, y) \\ & = \frac{1}{2} \sum_{x, y \in X} f^p(y)(f^p(y) - f^p(x))e^{\psi(y)}(e^{-\psi(y)} - e^{-\psi(x)})a(x, y) \\ & + \frac{1}{2} \sum_{x, y \in X} f^p(x)(f^p(y) - f^p(x))e^{-\psi(x)}(e^{\psi(x)} - e^{\psi(y)})a(x, y) \\ & \geq - \left( \sum_{x \in X} \Gamma(f^p, f^p)(x) \right)^{1/2} \left[ \left( \sum_{y \in X} f^{2p}(y)e^{2\psi(y)}\Gamma(e^{-\psi}, e^{-\psi})(y) \right)^{1/2} \right. \\ & \quad \left. + \left( \sum_{x \in X} f^{2p}(x)e^{-2\psi(x)}\Gamma(e^{\psi}, e^{\psi})(x) \right)^{1/2} \right]. \end{aligned} \quad (46)$$

For the second term in (45), we use Cauchy-Schwartz inequality to obtain

$$\begin{aligned} & \frac{1}{2} \sum_{x, y \in X} f^{2p}(x)(e^{-\psi(x)} - e^{-\psi(y)})(e^{\psi(x)} - e^{\psi(y)})a(x, y) = \sum_{x \in X} f^{2p}(x)\Gamma(e^{-\psi}, e^{\psi})(x) \\ & \geq - \left( \sum_{x \in X} f^{2p}(x)e^{2\psi(x)}\Gamma(e^{-\psi}, e^{-\psi})(x) \right)^{1/2} \left( \sum_{x \in X} f^{2p}(x)e^{-2\psi(x)}\Gamma(e^{\psi}, e^{\psi})(x) \right)^{1/2} \end{aligned} \quad (47)$$

For the last term in (45), we use Cauchy-Schwartz inequality to obtain

$$\begin{aligned} & \sum_{x, y \in X} f^{2p-1}(y)(f(x) - f(y))e^{\psi(y)}(e^{-\psi(y)} - e^{-\psi(x)})a(x, y) \\ & \geq -2 \left( \sum_{x \in X} f^{2p-2}(x)\Gamma(f, f)(x) \right)^{1/2} \left( \sum_{x \in X} f^{2p}(x)e^{2\psi(x)}\Gamma(e^{-\psi}, e^{-\psi})(x) \right)^{1/2}. \end{aligned} \quad (48)$$



We need two more elementary inequalities (Cf. (3.16) and (3.17) in [15]). For any non-negative function  $f \in L^2(X, \mu)$  and for all  $p \geq 1$ , we have

$$\sum_{x \in X} \Gamma(f^{2p-1}, f)(x) \geq \sum_{x \in X} f^{2p-2}(x) \Gamma(f, f)(x) \geq \frac{1}{2p-1} \sum_{x \in X} \Gamma(f^{2p-1}, f)(x) \quad (49)$$

and

$$\sum_{x \in X} \Gamma(f^p, f^p)(x) \geq \sum_{x \in X} \Gamma(f^{2p-1}, f)(x) \geq \frac{2p-1}{p^2} \sum_{x \in X} \Gamma(f^p, f^p)(x). \quad (50)$$

Since  $T$  is  $L$ -local, we use the inequality  $(e^a - 1)^2 \leq e^{2|a|} a^2$  to obtain

$$\max \left( e^{2\psi(x)} \Gamma(e^{-\psi}, e^{-\psi})(x), e^{-2\psi(x)} \Gamma(e^{\psi}, e^{\psi})(x) \right) \leq e^{2\text{osc}(\psi, L)} \Gamma(\psi, \psi)(x) \quad (51)$$

for all  $x \in X$ . Combining equations (43) through (51), we have

$$\begin{aligned} & \sum_{x \in X} \Gamma(e^{-\psi} f, e^{\psi} f^{2p-1})(x) \\ & \geq \frac{2p-1}{p^2} \sum_{x \in X} \Gamma(f^p, f^p)(x) - e^{2\text{osc}(\psi, L)} \sum_{x \in X} f^{2p}(x) \Gamma(\psi, \psi)(x) \\ & \quad - 4e^{\text{osc}(\psi, L)} \left( \sum_{x \in X} \Gamma(f^p, f^p)(x) \right)^{1/2} \left( \sum_{x \in X} f^{2p}(x) \Gamma(\psi, \psi)(x) \right)^{1/2}. \end{aligned}$$

We use the inequality  $4ab \leq a^2/(2p) + 8pb^2$  and  $p \geq 1$  to obtain

$$\begin{aligned} & \sum_{x \in X} \Gamma(e^{-\psi} f, e^{\psi} f^{2p-1})(x) \\ & \geq \frac{1}{2p} \sum_{x \in X} \Gamma(f^p, f^p)(x) - 9pe^{2\text{osc}(\psi, L)} \sum_{x \in X} f^{2p}(x) \Gamma(\psi, \psi)(x) \end{aligned}$$

for all bounded functions  $\psi$  and for all non-negative functions  $f \in L^2(X, \mu)$ .  $\square$

*Remark.* The elementary inequalities (49) and (50) are sometimes called Stroock-Varopolous inequalities. An unified approach to such inequalities is provided in [33, Lemma 2.4].

The next step is to bound the second term in the right side of (41) using cutoff Sobolev inequalities developed in Section 2.

**Lemma 3.5.** *Let  $(X, \mu)$  be a weighted graph satisfying (3) and satisfies  $\mathbf{V}(d_f)$  with volume growth exponent  $d_f$ . Let  $K$  be a  $\mu$ -symmetric Markov operator whose kernel  $k = k_1$  satisfies  $\mathbf{J}(\beta)$  for some  $\beta \geq 2$ . Further assume that  $K$  satisfies the cutoff Sobolev inequality  $\mathbf{CSJ}(\beta)$ . Let  $\Gamma$  denote the energy measure corresponding to  $K$  and let  $\Gamma_{(L)}$  correspond to the truncated version of  $\Gamma$  for some  $L > 0$ . Define*

$$\vartheta := \frac{1}{8(d_f + \beta)}.$$

Then there exists  $\lambda_0 \geq 1$ ,  $C_0 \geq 1$  such that the following property holds: For all  $\lambda \geq \lambda_0$ , for all  $p \geq 1$ , for all  $x \in X$ , for all  $r \geq 1$ , there exists a cut-off function  $\phi = \phi_{p,\lambda}$  for  $B(x, r) \subset B(x, 2r)$  such that for all non-negative functions  $f \in L^2(X, \mu)$  we have

$$\sum_{y \in X} \Gamma_{(\vartheta r)}(e^{-\lambda \phi} f, e^{\lambda \phi} f^{2p-1}) \geq \frac{1}{4p} \sum_{y \in X} \Gamma(f^p, f^p)(y) - \frac{C_0 p^\beta e^{8\beta \vartheta \lambda}}{r^\beta} \sum_{y \in X} f^{2p}(y) \mu(y), \quad (52)$$

where the cutoff function  $\phi$  above satisfies

$$\|\phi - \psi\|_\infty \leq \frac{1}{\lambda p}, \quad (53)$$

where  $\psi$  is given by (26).

*Proof.* By Lemma 3.4, for any cutoff function  $\phi$  for  $B(x, r) \subset B(x, 2r)$ , for all  $p \geq 1$ , for all  $\lambda > 0$ , and for all non-negative  $f \in L^2(X, \mu)$  we have

$$\begin{aligned} \sum_{y \in X} \Gamma_{(\vartheta r)}(e^{-\lambda \phi} f, e^{\lambda \phi} f^{2p-1}) \\ \geq \frac{1}{2p} \sum_{y \in X} \Gamma_{(\vartheta r)}(f^p, f^p)(y) - 9p\lambda^2 e^{2\lambda \text{osc}(\phi, L)} \sum_{y \in X} f^{2p}(y) \Gamma(\phi, \phi)(y). \end{aligned} \quad (54)$$

Using J( $\beta$ ) and a similar computation as [34, Equation (33)], there exists  $C_2 > 0$  such that

$$\sum_{y \in X} (\Gamma(f^p, f^p)(y) - \Gamma_{(\vartheta r)}(f^p, f^p)(y)) \leq C_2 r^{-\beta} \sum_{y \in X} f^{2p}(y) \mu(y) \quad (55)$$

for all  $r \geq 1$ , for all  $p \geq 1$  and for all  $f \in L^2(X, \mu)$ . By CSJ( $\beta$ ), there exists  $C_3, C_4 > 1$  such that for any  $n \in \mathbb{N}^*$  there exists a cutoff function  $\phi = \phi_n$  for  $B(x, r) \subset B(x, 2r)$  such that for all  $f \in L^2(X, \mu)$  with  $f \geq 0$ , we have

$$\sum_{y \in X} f^2(y) \Gamma(\phi, \phi)(y) \leq \frac{C_3}{n^2} \sum_{y \in X} \Gamma(f, f)(y) + \frac{C_4 \mathcal{G}_\beta(n)}{r^\beta} \sum_{y \in X} f^2(y) \mu(y), \quad (56)$$

where the function  $\phi = \phi_n$  above satisfies

$$\|\phi - \psi\|_\infty \leq n^{-1}. \quad (57)$$

We make the choice

$$n = \left\lceil 6p\lambda \exp(3\lambda\vartheta) \sqrt{C_3} \right\rceil \quad (58)$$

and  $\lambda_0 \geq 1$  such that  $n \geq \vartheta^{-1}$  for all  $\lambda \geq \lambda_0$  and for all  $p \geq 1$ . We will verify  $\phi$  satisfies the desired properties (52) and (53). Using (58) and (57), we immediately have (53). By (57), triangle inequality,  $n \geq \vartheta^{-1}$  and by the fact that  $\psi$  is  $r^{-1}$ -Lipschitz, we have

$$\text{osc}(\phi, \vartheta r) \leq \text{osc}(\psi, \vartheta r) + 2n^{-1} \leq \vartheta + 2\vartheta = 3\vartheta. \quad (59)$$

By (58), (56), (59), we have

$$\begin{aligned} & 9p\lambda^2 e^{2\lambda \operatorname{osc}(\phi, \vartheta r)} \sum_{y \in X} f^{2p}(y) \Gamma(\phi, \phi)(y) \\ & \leq \frac{1}{4p} \sum_{y \in X} \Gamma(f^p, f^p)(y) + \frac{9C_4 p \lambda^2 e^{6\vartheta \lambda} \mathcal{G}_\beta(n)}{r^\beta} \sum_{y \in X} f^{2p}(y) \mu(y), \end{aligned} \quad (60)$$

for all non-negative  $f \in L^2(X, \mu)$ . We use  $n \leq 12p\lambda e^{3\lambda\vartheta} \sqrt{C_3}$ ,  $\lambda \leq \vartheta^{-1} e^{\vartheta\lambda}$  and  $\mathcal{G}_\beta(n) \leq n^{\beta-1}$  for all  $n \in \mathbb{N}^*$  to obtain  $C_5 > 0$  (that depends on  $\vartheta, \beta$ ) such that

$$9C_4 p \lambda^2 e^{6\vartheta \lambda} \mathcal{G}_\beta(n) \leq 9C_4 p \lambda^2 e^{6\vartheta \lambda} (12p\lambda e^{3\lambda\vartheta} \sqrt{C_3})^{\beta-1} \leq C_5 p^\beta e^{4(\beta+1)\lambda\vartheta} \leq C_5 p^\beta e^{8\beta\lambda\vartheta} \quad (61)$$

for all  $p, n \geq 1$  and for all  $\lambda > \lambda_0$ . Combining (54), (55), (60) and (61), we obtain (52).  $\square$

We need the following Nash inequality to obtain off-diagonal estimates using Davies' method.

**Proposition 3.6** (Nash inequality). *Let  $(X, \mu)$  be a weighted graph satisfying (3) and satisfies  $\mathbf{V}(d_f)$  with volume growth exponent  $d_f$ . Let  $K$  be a  $\mu$ -symmetric Markov operator whose kernel  $k$  satisfies  $\mathbf{J}(\beta)$  for some  $\beta > 0$ . Let  $\mathcal{E}(\cdot, \cdot)$  denote the corresponding Dirichlet form. Then there exists  $C_N > 0$  such that*

$$\|f\|_2^{2(1+\beta/d_f)} \leq C_N \mathcal{E}(f, f) \|f\|_1^{2\beta/d_f} \quad (62)$$

for all  $f \in L^1(X, \mu)$ , where  $\|\cdot\|_p$  denotes the  $L^p(X, \mu)$  norm.

*Proof.* The proof of the Nash inequality (62) is essentially contained in [2, p. 1064]. We repeat the proof for convenience.

For  $r > 0$  and  $f \in L^1(X, \mu)$ , we denote by  $f_r : X \rightarrow \mathbb{R}$  the  $\mu$ -weighted average

$$f_r(x) = \frac{1}{V(x, r)} \sum_{y \in X} f(y) \mu(y).$$

We first bound  $\|f_r\|_1$ . There exists  $C_V > 0$  such that

$$\begin{aligned} \|f_r\|_1 &= \sum_{x \in X} |f_r(x)| \mu(x) \\ &\leq \sum_{x, y \in X} |f(y)| V(x, r)^{-1} \mu(y) \mu(x) = \sum_{y \in X} |f(y)| \mu(y) \sum_{x \in B(y, r)} V(x, r)^{-1} \mu(x) \\ &\leq C_V^2 \|f\|_1 \end{aligned} \quad (63)$$

for all  $f \in L^1(X, \mu)$  and for all  $r > 0$ . The second line above follows from triangle inequality and Fubini's theorem and the last line follows from  $\mathbf{V}(d_f)$ . By  $\mathbf{V}(d_f)$  and (3), there exists  $C_1 > 0$  such that

$$\|f_r\|_\infty \leq \|f\|_1 \sup_{x \in X} V(x, r)^{-1} \leq C_1 r^{-d_f} \|f\|_1 \quad (64)$$

$f \in L^1(X, \mu)$  and for all  $r > 0$ . Combining (63) and (64), there exists  $C_2 > 0$  such that

$$\|f_r\|_2 \leq \|f_r\|_1^{1/2} \|f_r\|_\infty^{1/2} \leq C_2 r^{-d_f/2} \|f\|_1 \quad (65)$$

for all  $f \in L^1(X, \mu)$  and for all  $r > 0$ .

There exists  $C_3 > 0$  such that for all  $f \in L^1(X, \mu)$  and for all  $r > 0$ , we have

$$\begin{aligned} \|f - f_r\|_2^2 &= \sum_{x \in X} |f(x) - f_r(x)|^2 \mu(x) \\ &\leq \sum_{x \in X} \frac{1}{V(x, r)} \sum_{y \in B(x, r)} (f(x) - f(y))^2 \mu(y) \mu(x) \\ &\leq C_1 \sum_{x, y \in X} (f(x) - f(y))^2 \mathbf{1}_{d(x, y) \leq r} r^{-d_f} \mu(x) \mu(y) \\ &\leq 2^{d_f + \beta} C_1 r^\beta \sum_{x, y \in X} (f(x) - f(y))^2 \frac{1}{(1 + d(x, y))^{d_f + \beta}} \mu(x) \mu(y) \\ &\leq \frac{C_3 r^\beta}{2} \sum_{x, y \in X} (f(x) - f(y))^2 k(x, y) \mu(x) \mu(y) = C_3 r^\beta \mathcal{E}(f, f). \end{aligned} \quad (66)$$

The second line above follows from Jensen's inequality, the third line follows from an application of  $\mathbf{V}(d_f)$  and (3) similar to (64) and the last line follows from  $\mathbf{J}(\beta)$  and Lemma 2.4(a).

By triangle inequality, (65) and (66), we have

$$\|f\|_2 \leq \|f_r\|_2 + \|f - f_r\|_2 \leq C_2 r^{-d_f/2} \|f\|_1 + C_3^{1/2} r^{\beta/2} \mathcal{E}(f, f)^{1/2} \quad (67)$$

for all  $f \in L^1(X, \mu)$  and all  $r > 0$ . The Nash inequality (62) follows from (67) and the choice

$$r = \left( \frac{\|f\|_1^2}{\mathcal{E}(f, f)} \right)^{1/(d_f + \beta)}.$$

□

The following lemma is analogous to [15, Lemma 3.21] but the statement and its proof is slightly modified to suit our context.

**Lemma 3.7.** *Let  $w : [0, \infty) \rightarrow (0, \infty)$  be a non-decreasing function and suppose that  $u \in C^1([0, \infty); (0, \infty))$  satisfies*

$$u'(t) \leq -\frac{\epsilon}{p} \left( \frac{t^{(p-2)/\theta p}}{w(t)} \right)^{\theta p} u^{1+\theta p}(t) + \delta p^\beta u(t) \quad (68)$$

for some positive  $\epsilon, \theta$  and  $\delta, \beta \in [2, \infty)$  and  $p = 2^k$  for some  $k \in \mathbb{N}^*$ . Then  $u$  satisfies

$$u(t) \leq \left( \frac{2p^{\beta+1}}{\epsilon \theta} \right)^{1/\theta p} t^{(1-p)/\theta p} w(t) e^{\delta t/p} \quad (69)$$

*Proof.* Set  $v(t) = e^{-\delta p^\beta t} u(t)$ . By (68), we have

$$v'(t) = e^{-\delta p^\beta t} \left( u(t) - \delta p^\beta u(t) \right) \leq -\frac{\epsilon t^{p-2}}{p w(t)^{\theta p}} e^{\theta \delta p^{\beta+1} t} v(t)^{1+\theta p}.$$

Hence

$$\frac{d}{dt} (v(t))^{-\theta p} \geq \epsilon \theta t^{p-2} w(t)^{-\theta p} e^{\theta \delta p^{\beta+1} t}$$

and so, since  $w$  is non-decreasing

$$e^{\delta \theta p^{\beta+1} t} u(t)^{-\theta p} \geq \epsilon \theta w(t)^{-\theta p} \int_0^t s^{(p-2)} e^{\theta \delta p^{\beta+1} s} ds. \quad (70)$$

Note that

$$\begin{aligned} \int_0^t s^{(p-2)} e^{\theta \delta p^{\beta+1} s} ds &\geq (t/\delta \theta p^{\beta+1})^{p-1} \int_{\delta \theta p^{\beta+1}(1-1/p^{\beta+1})}^{\delta \theta p^{\beta+1}} y^{(p-2)} e^{ty} dy \\ &\geq \frac{t^{p-1}}{p-1} \exp\left(\delta \theta p^{\beta+1} t - \delta \theta t\right) \left[1 - (1 - p^{-\beta-1})^{p-1}\right] \\ &\geq \frac{t^{p-1}}{2p^{\beta+1}} \exp\left(\delta \theta p^{\beta+1} t - \delta \theta t\right). \end{aligned} \quad (71)$$

In the last line above, we used the bound  $(1 - p^{-\beta-1})^{p-1} \geq 1 - p^{-\beta-1}(p-1)$  for all  $p, \beta \geq 2$ . Combining (70) and (71) yields (69).  $\square$

We now have all the necessary ingredients to prove Proposition 3.1.

*Proof of Proposition 3.1.* Let  $H_t := \exp(t(K - I))$  denote the corresponding continuous time semigroup. Note that  $H_t$  is  $\mu$ -symmetric and let  $h_t$  be the kernel of  $H_t$  with respect to  $\mu$ . As explained in [34, Remark 3], by [22, Theorem 3.6] it suffices to show the following bound on  $h_t$ : there exists  $C_1 > 0$  such that

$$h_t(x, y) \leq C_1 \min\left(\frac{1}{t^{d_f/\beta}}, \frac{t}{(d(x, y))^{d_f+\beta}}\right) \quad (72)$$

for all  $x, y \in X$  and for all  $t \geq 1$ . By [18, Proposition II.1] the Nash inequality (62) in Proposition 3.6 implies that there exists  $C_2 > 0$  such that

$$h_t(x, y) \leq \frac{C_2}{t^{d_f/\beta}} \quad (73)$$

for all  $x, y \in X$  and for all  $t \geq 1$ .

Let  $t \geq 1$  and let  $x, y \in X$  such that  $d(x, y) \geq 2$ . Let  $\vartheta = 1/(8(d_f + \beta))$  and  $\lambda_0, C_0 \geq 1$  be constants given by Lemma 3.5. We define parameters  $L$  and  $r$  that depend only on  $x$  and  $y$  as

$$r = d(x, y)/2, L = \vartheta r. \quad (74)$$

Let  $\lambda \geq \lambda_0$  be arbitrary. Let  $\Gamma$  denote the energy measure corresponding to  $K$  and let  $\Gamma_{(L)}$  denote its truncation. Define the truncated Dirichlet form

$$\mathcal{E}_L(f, f) = \sum_{y \in X} \Gamma_{(L)}(f, f)(y)$$

for all  $f \in L^2(X, \mu)$ . Let  $h_t^{(L)}$  denote the continuous time kernel with respect to  $\mu$  for the corresponding jump process and let  $H_t^{(L)}$  denote the corresponding Markov semigroup. Let  $p_k = 2^k$  for  $k \in \mathbb{N}$  and let  $\psi_k = \lambda \phi_{p_k, \lambda}$  where  $\phi_{p_k, \lambda}$  is the cutoff function for  $B(x, r) \subset B(x, 2r)$  given by Lemma 3.5. Define the ‘perturbed semigroup’

$$H_t^{L, \psi_k} f(x) = e^{\psi_k(x)} \left( H_t^{(L)} \left( e^{-\psi_k} f \right) \right) (x). \quad (75)$$

We now pick  $f \in L^2(M, \mu)$  and  $f \geq 0$  with  $\|f\|_2 = 1$  and define

$$f_{t,k} := H_t^{L, \psi_k} f \quad (76)$$

for all  $k \in \mathbb{N}$ , and  $P_t^{\psi_k}$  denotes the perturbed semigroup defined in (75). We remark that the constants below do not depend on the choice of  $x, y \in X$  with  $d(x, y) \geq 2$  or the choice of  $f \in L^2(M, \mu)$ ,  $t \geq 1$  or  $\lambda \geq \lambda_0$ .

Using Lemma 3.5, we have

$$\begin{aligned} \frac{d}{dt} \|f_{t,0}\|_2^2 &= -2\mathcal{E} \left( e^{\psi_1} f_{t,0}, e^{-\psi_1} f_{t,0} \right) \\ &\leq 2C_0 \frac{e^{8\beta\vartheta\lambda}}{r^\beta} \|f_{t,0}\|_2^2 \end{aligned} \quad (77)$$

and

$$\begin{aligned} \frac{d}{dt} \|f_{t,k}\|_{2p_k}^{2p_k} &= -2p_k \mathcal{E} \left( e^{\psi_k} f_{t,k}^{2p_k-1}, e^{-\psi_k} f_{t,k} \right) \\ &\leq -\frac{1}{2} \mathcal{E} \left( f_{t,k}^{p_k}, f_{t,k}^{p_k} \right) + 2C_0 \frac{e^{8\beta\vartheta\lambda} p_k^{\beta+1}}{r^\beta} \|f_{t,k}\|_{2p_k}^{2p_k} \end{aligned} \quad (78)$$

for all  $k \in \mathbb{N}^*$ . By solving (77), we obtain

$$\|f_{t,0}\|_{p_1} = \|f_{t,0}\|_2 \leq \exp \left( C_0 e^{8\beta\vartheta\lambda} t / r^\beta \right) \|f\|_2 = \exp \left( C_0 e^{8\beta\vartheta\lambda} t / r^\beta \right). \quad (79)$$

Using (78) and Nash inequality (62), we obtain

$$\frac{d}{dt} \|f_{t,k}\|_{2p_k} \leq -\frac{1}{4C_N p_k} \|f_{t,k}\|_{2p_k}^{1+2\beta p_k/d_f} \|f_{t,k}\|_{p_k}^{-2\beta p_k/d_f} + C_0 p_k^\beta \frac{e^{8\beta\vartheta\lambda}}{r^\beta} \|f_{t,k}\|_{2p_k} \quad (80)$$

for all  $k \in \mathbb{N}^*$ . By (53), we have  $\lambda \|\psi_k - \psi_{k+1}\|_\infty \leq 3/(2p_k)$ . This along the fact that  $H_t^{(L)}$  is a contraction on  $L^\infty$  yields

$$\exp(-3/p_k) f_{t,k+1} \leq f_{t,k} \leq \exp(3/p_k) f_{t,k+1} \quad (81)$$

for all  $k \in \mathbb{N}_{\geq 0}$ . Combining (80) and (81), we obtain

$$\frac{d}{dt} \|f_{t,k}\|_{2p_k} \leq -\frac{1}{C_A p_k} \|f_{t,k}\|_{2p_k}^{1+2\beta p_k/d_f} \|f_{t,k-1}\|_{p_k}^{-2\beta p_k/d_f} + C_0 p_k^\beta \frac{e^{8\beta\vartheta\lambda}}{r^\beta} \|f_{t,k}\|_{2p_k} \quad (82)$$

for all  $k \in \mathbb{N}^*$ , where  $C_A = 4C_N \exp(12\beta/d_f)$ .

Let  $u_k(t) = \|f_{t,k-1}\|_{p_k}$  and let

$$w_k(t) = \sup \{ s^{d_f(p_k-2)/(2\beta p_k)} u_k(s) : s \in (0, t] \}.$$

By (79),  $w_1(t) \leq \exp(C_0 e^{8\beta\vartheta\lambda} t/r^\beta)$ . Further by (82),  $u_{k+1}$  satisfies (68) with  $\epsilon = 1/C_A$ ,  $\theta = 2\beta/d_f$ ,  $\delta = C_0(e^{8\beta\vartheta\lambda}/r^\beta)$ ,  $w = w_k$ ,  $p = p_k$ . Hence by Lemma 3.7,

$$u_{k+1}(t) \leq (2^{(\beta+1)k+1}/\epsilon\theta)^{1/(\theta p_k)} t^{(1-p_k)/\theta p_k} e^{\delta t/p_k} w_k(t).$$

Therefore

$$w_{k+1}(t)/w_k(t) \leq (2^{(\beta+1)k+1}/\epsilon\theta)^{1/(\theta 2^k)} e^{\delta t/2^k}$$

for  $k \in \mathbb{N}^*$ . Hence, we obtain

$$\lim_{k \rightarrow \infty} w_k(t) \leq C_3 e^{\delta t} w_1(t) \leq C_2 \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta)$$

where  $C_3 = C_3(\beta, \epsilon, \theta) > 0$ . Hence

$$\lim_{k \rightarrow \infty} u_k(t) = \|H_t^{L, \psi_\infty} f\|_\infty \leq \frac{C_2}{t^{d_f/2\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta).$$

where  $\psi_\infty = \lim_{k \rightarrow \infty} \psi_k$ . Since the above bound holds for all non-negative  $f \in L^2(X, \mu)$  we obtain

$$\|H_t^{L, \psi_\infty}\|_{2 \rightarrow \infty} \leq \frac{C_3}{t^{d_f/2\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta).$$

The estimate is unchanged if we replace  $\psi_k$ 's by  $-\psi_k$ . Since  $H_t^{L, -\psi}$  is the adjoint of  $H_t^{L, \psi}$ , by duality we have that for all  $t > 0$

$$\|H_t^{L, \psi_\infty}\|_{1 \rightarrow 2} \leq \frac{C_3}{t^{d_f/2\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta).$$

Combining the above, we have

$$\|H_t^{L, \psi_\infty}\|_{1 \rightarrow \infty} \leq \frac{C_3 2^{d_f/\beta}}{t^{d_f/\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta).$$

for all  $\lambda \geq \lambda_0$ , for all  $x, y \in M$  with  $d(x, y) \geq 2$ , where  $L$  is as defined in (84). Therefore

$$h_t^{(L)}(x, y) \leq \frac{C_2 2^{d_f/\beta}}{t^{d_f/\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta + \psi_\infty(y) - \psi_\infty(x)). \quad (83)$$

for all  $x, y \in X$  satisfying  $d(x, y) \geq 2$ , for all  $t \geq 1$  and for all  $\lambda \geq \lambda_0$ .

Now we choose  $\lambda$  as

$$\lambda = \frac{d_f + \beta}{\beta} \log(r^\beta/t). \quad (84)$$

By (73), it suffices to show (72) for the case  $d(x, y) \geq 2$  and  $\lambda \geq \lambda_0$ . By the choices in (74), we have  $\psi_\infty(y) - \psi_\infty(x) = -\lambda$  and hence by (83), we obtain

$$h_t^{(L)}(x, y) \leq \frac{C_2 2^{d_f/\beta}}{t^{d_f/\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta - \lambda) = \frac{C_2 2^{d_f/\beta}}{t^{d_f/\beta}} \exp(2C_0) \frac{t^{1+(d_f/\beta)}}{r^{d_f+\beta}}. \quad (85)$$

By (85), (74) along with Meyer decomposition bound in [11, Lemma 3.1], there exists  $C_4, C_5 > 0$  such that for all  $x, y \in X$  and for all  $t \geq 1$  satisfying  $d(x, y)^\beta \geq C_5 t$ , we have

$$h_t(x, y) \leq \frac{C_4 t}{(d(x, y))^{d_f+\beta}}.$$

The above equation along with (73) yields (72) which in turn implies  $\text{UHKP}(d_f, \beta)$ .  $\square$

We will now prove the main results stated in the introduction.

*Proof of Theorem 1.6.* Proposition 3.1 and Proposition 2.1 are the two desired implications.  $\square$

*Proof of Theorem 1.5.* Let  $K_1$  satisfy  $\text{HKP}(d_f, \beta)$ . Then by Theorem 1.5, we have that  $K_1$  satisfies the cutoff Sobolev inequality  $\text{CSJ}(\beta)$ . Since two  $\mu$ -symmetric Markov operators satisfying  $\text{J}(\beta)$  have comparable energy measures,  $K_2$  satisfies  $\text{CSJ}(\beta)$ . Therefore by Theorem 1.6,  $K_2$  satisfies  $\text{UHKP}(d_f, \beta)$ . The upper bounds  $\text{UHKP}(d_f, \beta)$  and (72) are sufficient to show the matching lower bounds using an iteration argument due to Bass and Levin [14]. The argument in [34] can be directly adapted to this setting. Therefore  $K_2$  satisfies  $\text{HKP}(d_f, \beta)$ .  $\square$

**Acknowledgement.** We thank Martin Barlow for inspiring conversations and for his helpful suggestions on an earlier draft. We thank Tom Hutchcroft for proofreading part of the manuscript.

## REFERENCES

- [1] S. Andres, M. T. Barlow, *Energy inequalities for cutoff-functions and some applications*, J. Reine Angew. Math. 699 (2015), 183–215. [MR3305925](#)
- [2] Bakry, D.; Coulhon, T.; Ledoux, M; Saloff-Coste, L. Sobolev Inequalities in Disguise, *Indiana Univ. Math. J.* **44** (1995), no. 4, 1033–1074. [MR1386760](#)
- [3] M. T. Barlow, *Diffusions on fractals*, Lectures on probability theory and statistics (Saint-Flour, 1995), 1–121, Lecture Notes in Math., 1690, Springer, Berlin, 1998. [MR1668115](#)
- [4] M. T. Barlow, *Which values of the volume growth and escape time exponent are possible for a graph?*, Rev. Mat. Iberoamericana 20 (2004), no. 1, 1–31. [MR2076770](#)
- [5] M. T. Barlow, R. F. Bass, *Brownian motion and harmonic analysis on Sierpinski carpets*, Canad. J. Math. 51 (1999), no. 4, 673–744. [MR1701339](#)
- [6] M. T. Barlow, R. F. Bass, *Random walks on graphical Sierpinski carpets*, Random walks and discrete potential theory (Cortona, 1997), 26–55, Sympos. Math., XXXIX, Cambridge Univ. Press, Cambridge, 1999. [MR1802425](#)
- [7] M. T. Barlow, R. F. Bass, *Stability of parabolic Harnack inequalities*, Trans. Amer. Math. Soc. 356 (2004), no. 4, 1501–1533. [MR2034316](#)
- [8] M. T. Barlow, R. F. Bass, T. Kumagai, *Stability of parabolic Harnack inequalities on metric measure spaces*, J. Math. Soc. Japan 58 (2006), no. 2, 485–519. [MR2228569](#)
- [9] M. T. Barlow, R. F. Bass, T. Kumagai, *Parabolic Harnack inequality and heat kernel estimates for random walks with long range jumps*, Math. Z. 261 (2009), 297–320. [MR2457301](#)
- [10] M. T. Barlow, T. Coulhon, T. Kumagai, *Characterization of sub-Gaussian heat kernel estimates on strongly recurrent graphs*, Comm. Pure Appl. Math. 58 (2005), no. 12, 1642–1677. [MR2177164](#)
- [11] M. T. Barlow, A. Grigor’yan, T. Kumagai, *Heat kernel upper bounds for jump processes and the first exit time*, J. Reine Angew. Math., 626 (2009) 135–157. [MR2492992](#)
- [12] M. T. Barlow, E. A. Perkins, *Brownian motion on the Sierpinski gasket*, Probab. Theory Related Fields 79 (1988), no. 4, 543–623. [MR0966175](#)
- [13] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular variation*, Encyclopedia of Mathematics and its Applications, vol. 27, Cambridge University Press, Cambridge, 1987. [MR898871](#)



- [14] R. F. Bass, D. A. Levin, *Transition probabilities for symmetric jump processes*, Trans. Amer. Math. Soc. 354, no. 7 (2002), 2933–2953 [MR1895210](#)
- [15] E. A. Carlen, S. Kusuoka, D. W. Stroock, *Upper bounds for symmetric Markov transition functions*, Ann. Inst. Henri Poincaré - Probab. Statist., 23 (1987), 245–287. [MR0898496](#)
- [16] Z.-Q. Chen, T. Kumagai, *Heat kernel estimates for stable-like processes on  $d$ -sets*, Stochastic Process Appl. 108 (2003), 27–62. [MR2008600](#)
- [17] Z.-Q. Chen, T. Kumagai, *Heat kernel estimates for jump processes of mixed types on metric measure spaces*, Stochastic Process Appl. 108 (2003), 27–62. [MR2357678](#)
- [18] T. Coulhon, *Ultracontractivity and Nash Type Inequalities* Jour. of Func. Anal., 141, 510–539 (1996) [MR1418518](#)
- [19] T. Coulhon, A. Grigor'yan, *On-diagonal lower bounds for heat kernels and Markov chains*, Duke Math. J. 89 (1997), no. 1, 133–199. [MR1458975](#)
- [20] E. B. Davies, *Explicit constants for Gaussian upper bounds on heat kernels*, Amer. J. Math. 109 (1987), no. 2, 319–333. [MR0882426](#)
- [21] E. B. Davies, *Uniformly elliptic operators with measurable coefficients*, J. Funct. Anal. 132 (1995), no. 1, 141–169. [MR1346221](#)
- [22] T. Delmotte, *Parabolic Harnack inequality and estimates of Markov chains on graphs*, Revista Mat. Iberoamericana, 15, (1999). [MR1681641](#)
- [23] A. Grigor'yan, *Heat kernels and function theory on metric measure spaces*, Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), 143–172, Contemp. Math., 338, Amer. Math. Soc., Providence, RI, 2003. [MR2039954](#)
- [24] A. Grigor'yan, J. Hu, K.-S. Lau, *Estimates of heat kernels for non-local regular Dirichlet forms*, Trans. Amer. Math. Soc. 366 (2014), no. 12, 6397–6441. [MR3267014](#)
- [25] A. Grigor'yan, T. Kumagai, *On the dichotomy in the heat kernel two sided estimates*, Analysis on graphs and its applications, 199–210, Proc. Sympos. Pure Math., 77, Amer. Math. Soc., Providence, RI, 2008. [MR2459870](#)
- [26] A. Grigor'yan, A. Telcs, *Sub-Gaussian estimates of heat kernels on infinite graphs*, Duke Math. J. 109 (2001), no. 3, 451–510. [MR1853353](#)
- [27] A. Grigor'yan, A. Telcs, *Harnack inequalities and sub-Gaussian estimates for random walks*, Math. Ann. 324, 551–556 (2002). [MR1938457](#)
- [28] B. M. Hambly, T. Kumagai, *Transition density estimates for diffusion processes on post critically finite self-similar fractals*, Proc. London Math. Soc. (3) 78 (1999), no. 2, 431–458. [MR1665249](#)
- [29] O. D. Jones, *Transition probabilities for the simple random walk on the Sierpinski graph*, Stochastic Process. Appl. 61 (1996), no. 1, 45–69. [MR1378848](#)
- [30] J. Kigami, *Analysis on fractals*, Cambridge Tracts in Mathematics, 143. Cambridge University Press, Cambridge, 2001. viii+226 pp. ISBN: 0-521-79321-1 [MR1840042](#)
- [31] T. Kumagai, *Some remarks for stable-like jump processes on fractals*, Fractals in Graz 2001, 185–196, Trends Math., Birkhäuser, Basel, 2003. [MR2091704](#)
- [32] P.-A. Meyer, *Renaissance, recollements, mélanges, ralentissement de processus de Markov*, Ann. Inst. Fourier 25 (3-4) (1975), 464–497. [MR0415784](#)
- [33] E. Mossel, K. Oleszkiewicz, A. Sen, *On reverse hypercontractivity*, Geom. Funct. Anal. 23 (2013), no. 3, 1062–1097. [MR3061780](#)
- [34] M. Murugan, L. Saloff-Coste, *Transition probability estimates for long range random walks*, New York J. Math. 21 (2015), 723–757. [MR3386544](#)
- [35] M. Murugan, L. Saloff-Coste, *Anomalous threshold behavior of long range random walks*, Electron. J. Probab. 20 (2015), no. 74, 21 pp. [MR3371433](#)
- [36] M. Murugan, L. Saloff-Coste, *Davies' method for anomalous diffusions*, in preparation.
- [37] A. Stós, *Symmetric  $\alpha$ -stable processes on  $d$ -sets*, Bull. Polish Acad. Sci. Math. 48 (2000), no. 3, 237–245. [MR1779007](#)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA AND PACIFIC INSTITUTE  
FOR THE MATHEMATICAL SCIENCES, VANCOUVER, BC V6T 1Z2, CANADA.

*E-mail address*, M. Murugan: `mathav@math.ubc.ca`

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NY 14853, USA.

*E-mail address*, L. Saloff-Coste: `lsc@math.cornell.edu`